ON PARAMETERIZING HETEROGENEITY AND INCORPORATING GEOPHYSICAL MEASUREMENTS IN HYDROGEOPHYSICAL INVERSE MODELING

Michael B. Kowalsky, Michael Commer, Kenneth H. Williams, and Stefan Finsterle

Lawrence Berkeley National Laboratory
Berkeley, California, 94720, U.S.A.
e-mail: mbkowalsky@lbl.gov

ABSTRACT

We consider the estimation of hydrological parameters through inverse modeling of hydrological (tracer) data and/or geophysical (electrical resistivity) data associated with experiments from the shallow unconfined uranium-contaminated aquifer at the DOE Integrated Field Research Challenge (IFRC) site at Rifle, Colorado. The purpose of this study, which draws on examples from recent and ongoing research, is to highlight the degree to which certain modeling decisions impact inversion results. Through synthetic examples based on field experiments with real-world complexities, we focus on decisions related to (1) how heterogeneity and other hydrological features are parameterized, and (2) how geophysical monitoring data are incorporated.

INTRODUCTION

Obtaining estimates of hydrological properties through inverse modeling requires decisions of great consequence to be made on how to parameterize heterogeneity (i.e., how to represent a heterogeneous distribution using a limited number of parameters that are amenable to estimation) and how to include different types of measurements in a model. Inadequate parameterization of heterogeneity or other hydrological features leads to errors in the model structure that are partly compensated for by biased property estimates, which may allow for an improved fit to the calibration data but lead to incorrect interpretations of hydrological phenomena and reduce the ability of the model to make reliable predictions. The manner in which observations are integrated in inverse modeling is similarly important. If the procedure for simulating measurements is not consistent with the true measurement process, then the resulting property estimates may be biased.

The purpose of this study is to highlight the degree to which certain modeling decisions can affect inverse modeling results. Through synthetic examples based on field experiments with real-world complexities, we focus on two areas in particular: (1) how heterogeneity and other hydrological features are parameterized, and (2) how geophysical monitoring data are incorporated.

An overview of the methodology and the field experiments that provide motivation for the study are given, followed by some examples and conclusions.

METHODOLOGY

The inverse modeling approach, as described by Kowalsky et al. (2012), consists of three main parts: a forward model, parameterization of heterogeneity and hydrological features, and an inverse modeling framework that integrates different types of data. In this study, TOUGH2 (Pruess et al., 1999) provides the forward model used to simulate the flow and transport phenomena in the experiment and the corresponding measurements for a given set of input parameters.

Simplifying assumptions made in the parameterization of heterogeneity must be carefully considered to insure they are adequate for a given application and consistent with site conditions (e.g., Moore and Doherty, 2006). A variety of approaches are available. A pixel-based approach divides the model into discrete regions, often in a regularly spaced grid or layers, allowing for values at each pixel to be estimated through inversion. In the related zonation approach, uniform properties are assigned to model regions based on practical considerations or based on spatial information related to the geology that is available from characterization data (e.g., geophysical data or core descriptions) to a varying degree of spatial coverage and accuracy. A geostatistical parameterization treats the heterogeneous property as a spatially correlated random field that is a function of the semivariogram and conditioning values at so-called pilot point locations estimated during inversion. Various aspects of these parameterization approaches or combinations thereof are considered in the examples given below.

The inverse modeling framework used here is iTOUGH2 (Finsterle, 2004). It provides capabilities for parameter estimation, error analysis and uncertainty propagation for TOUGH2 and other linked forward models, including geophysical models. Suitable geophysical measurements are made a function of properties that are simulated in the TOUGH2 forward model and, therefore, a function of the hydrological input parameters. In other words, iTOUGH2 allows for coupled inversion of hydrological and geophysical data (or hydrogeophysical inversion). A variety of geophysical measurements have been integrated into iTOUGH inversions, including time-lapse ground-penetrating radar, electrical resistivity, and seismic (Kowalsky et al., 2004, 2005, 2008, 2010, 2011).

Geophysical monitoring data can be included in hydrological inversions in different ways (e.g., Hinnell et al., 2010). In one approach, the geophysical observations are directly simulated based on TOUGH2 state variables at a given time and compared to the measured values as part of the calibration procedure. A petrophysical model converts the TOUGH2 state variables to the input needed for the geophysical simulation. For example, Archie's law (1942) relates porosity, saturation, and fluid electrical conductivity—which itself is a function of the concentration of dissolved mass components—to bulk electrical resistivity, the input property needed for an electrical resistivity simulation. Commer et al. (2012) present an example, which is also related to this study, using this approach.

In a second approach, traditional geophysical tomography is performed as a pre-processing step. The resulting tomogram of the relevant geophysical property is then used as spatially distributed calibration data in the inversion, typically also requiring a petrophysical model to relate the simulated hydrological response to the geophysical tomogram. While the sequential application of geophysical tomography and hydrological inversion is a straightforward means for incorporating geophysical data, errors that are commonly introduced in the tomography procedure (such as smoothing or streaks in the image) may cause the hydrological parameter estimates to be biased, as is demonstrated in an example below.

OVERVIEW OF FIELD EXPERIMENTS

The examples in this study were motivated by experiments at the U.S. Department of Energy (DOE) Integrated Field Research Challenge Site (IFRC) at Rifle, Colorado in a shallow unconfined aquifer that was contaminated from uranium mill tailings. The aquifer is 2.5 to 3 m thick, and is located on the floodplain of the Colorado River in alluvium situated above an impermeable bedrock formation. A main objective at the IFRC site has been to perform biostimulation experiments for uranium remediation through the injection of acetate. Nonreactive (i.e., conservative) tracers are typically also injected into the groundwater to aid in characterization of the transport processes and properties of the subsurface. Detailed descriptions of the IFRC site and related field experiments are given elsewhere (Anderson et al., 2003; Vrionis et al., 2005; Yabusaki et al., 2007; Williams et al., 2011). In the examples given below, we consider the "Winchester" experiment of 2007 (Williams et al., 2011), and the "Best Western" experiment of 2011.

Understanding subsurface heterogeneity is key to analyzing complex biogeochemical reactions such as those occurring in these field experiments. Heterogeneity is expected to have a large impact on the efficacy of biostimulation. Thus, developing, testing, and analyzing techniques for estimating heterogeneous properties based on limited characterization data, such as tracer concentrations, core data, and geophysical data, continues to be a critical research topic in hydrogeology.

EXAMPLE 1. PARAMETERIZING HETEROGENEITY

A recent study by Kowalsky et al. (2012) on inverse modeling of synthetic and actual field data from the Winchester experiment highlights some issues related to parameterization, and forms the basis for discussion in this example. The experiment involved the injection of groundwater amended with acetate, to induce biostimulation, and sodium bromide, to act as a conservative tracer and allow for investigation of hydrological properties at the site. The amended groundwater was injected into 10 wells (shown with triangles in Figure 1) situated perpendicularly to the direction of groundwater flow, and the time-varying concentrations were measured in 12 down-gradient monitoring wells (shown with circles in Figure 1).

To investigate how various choices related to the parameterization of heterogeneity in such a system can affect inverse modeling results, a hydrological model was developed to simulate the experiment and the measurements. Based on a permeability distribution generated in GSLIB (Deutsch and Journel, 1992) using Gaussian sequential simulation (SGSIM), synthetic bromide concentration data were simulated for different sampling intervals and noise levels. Figure 2 shows a best-case scenario data set with high temporal sampling and no added noise. The permeability distribution, as shown in Figure 1a, represents the true model for the synthetic study. Snapshots of the concentration at several times are shown in Figure 3.

When characterization data related to the spatial distribution of geological properties or facies are available (e.g., from geophysical data, well logging data or core descriptions), it is naturally of interest to incorporate this information into a hydrological model. For this purpose, a zonation parameterization may be used in which each zone or facies is modeled with uniform properties that are measured somehow or estimated through inversion. To explore some aspects of this approach, several facies "data sets" were derived from the true permeability distribution—with varying resolution, coverage, and accuracy—and used for inversion.

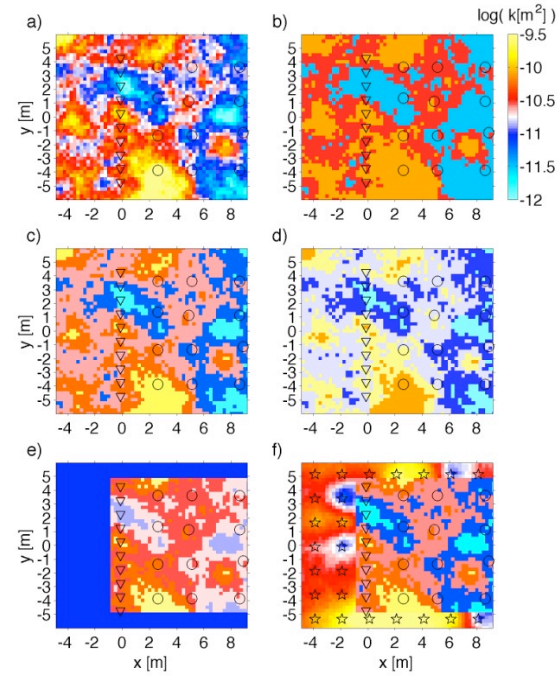

Figure 1. Inversion of synthetic data with zones or "facies" information of varying accuracy and completeness. (a) True permeability. Inversion with the geometry of (b) three zones known perfectly, (c) five zones known perfectly, (d) slightly inaccurately, and (e) perfectly but only for the well region with core data. (f) Inversion with geometry of five zones known in the well region, and pilot points used elsewhere.

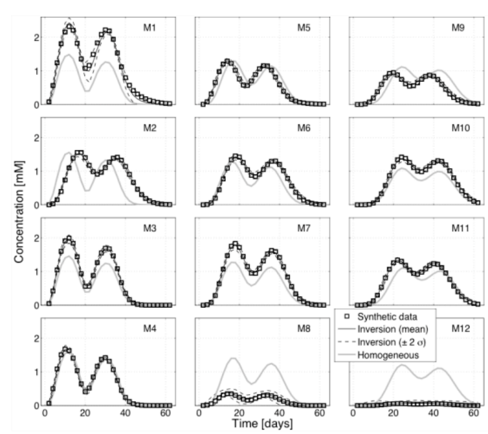

Figure 2. Noise-free bromide concentrations used as synthetic data (symbols), fit by inversion with a homogeneous model (solid gray line) and by inversion with a geostatistical parameterization (black lines show mean with +/- 2 standard deviations).

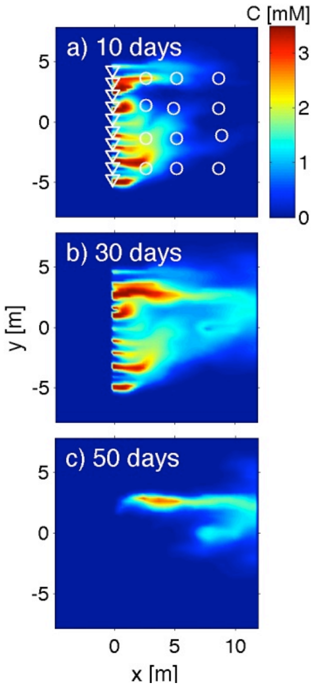

Figure 3. Bromide concentration at several times.

For the case in which facies data comprise 3 zones (permeability ranges), the inversion is not able to obtain a good fit between the simulated and measured (synthetic) data, and the resulting parameter estimates are inaccurate (Figure 1b), revealing an inadequate parameterization. On the other hand, when the facies data comprise 5 zones, inversion provides a nearly perfect data fit and accurate parameter estimates (Figure 1c). Introducing a small amount of error in the facies geometry, by introducing 50 cm shifts in the geometry in 2 m by 2 m regions, considerably worsens the parameter estimates (Figure 1d). When the facies data are without error but limited to the region near the wells where core data are available, then the parameter estimates are again poor (Figure 1e). However, combining the facies data from the limited region with a geostatistical parameterization, in which pilot point values are also estimated in the inversion, allows for accurate parameter estimates (Figure 1f). Thus, accounting for errors in the zonal geometry and properly parameterizing regions where no information is available (e.g., up-gradient of the injection wells) is clearly important for accurate inverse modeling results.

Implementing pilot points in a geostatistical parameterization must also be done with care. Rather than placing pilot points throughout the entire model domain, a practitioner might choose to reduce the number of unknowns by modeling the region up-gradient (to the left) of the wells using a uniform zone with a single unknown parameter. However, the estimated permeability distribution is then inaccurate (Figure 4a) and does not allow for an adequate fit to the data. When the region is modeled with unconditional heterogeneity (i.e., reflecting the correct spatial correlation structure but not controlled by pilot points), the data fit and the parameter estimates are similarly not optimal (Figure 4c). The best inversion results are achieved by also including pilot points in the up-gradient region (Figure 4e).

Kowalsky et al. (2012) provide metrics to further quantify improvement between cases. Deteriorating experiment conditions, variable porosity, and errors in the semivariogram and gradient direction are also considered, as is application to the field data and a 3D model.

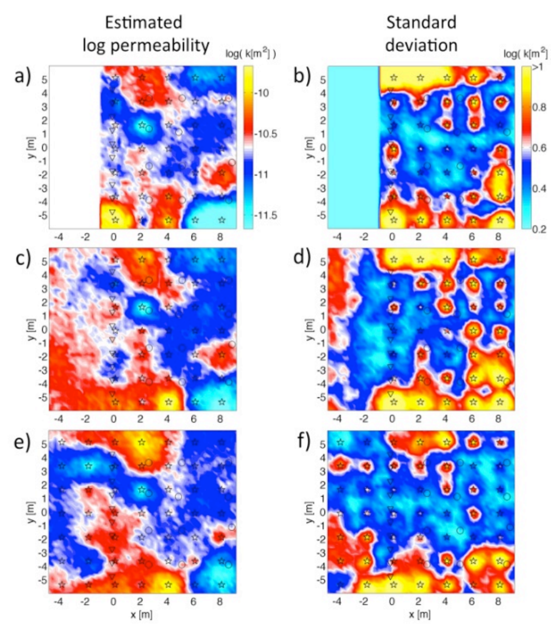

Figure 4. Sensitivity of geostatistical inversion to the parameterization in the region up-gradient (to the left) of the injection and monitoring wells. The region is parameterized with (a) one zone of uniform permeability, (b) unconditional heterogeneity (no pilot points), and (c) two columns of pilot points. Figure 1a shows the true permeability model.

EXAMPLE 2. INCORPORATING GEOPHYSICAL MONITORING DATA

This second example is motivated by the Best Western experiment which sought to study, among other things, the impact of prolonged bicarbonate injection on metal oxyanion (U, As, V, Se, Mo) mobility. In one portion of the area used for this experiment (plot C at the IFRC site), sodium bicarbonate was injected into the groundwater, having the added benefit of serving as an electrically conductive tracer (while not strictly conservative, it can be considered as a tracer to a first approximation). Electrical conductivity (EC) measurements were regularly made on fluid samples from the monitoring wells. In addition, crosshole electrical resistance tomography (ERT) data were collected at a number of times.

A hydrological model, representing a 2D vertical plane aligned with the direction of groundwater flow, was constructed for the vicinity of the bicarbonate injection (Figure 5). The model includes the vadose zone in the upper 3 meters and a saturated zone in the lower 3 meters (only the saturated zone is depicted). The top boundary is given by constant atmospheric pressure, while the bottom is a no-flow boundary. A gradient is added to the constant hydrostatic pressure specified on the side boundaries to induce lateral groundwater flow. The amended groundwater is injected into a high-permeability well at x = 3 m.

A hypothetical model is considered with simple layered heterogeneity given by a zone of higher log permeability (-10.5) between 4.2 and 5.3 m depth and a moderately lower log permeability (-11) elsewhere. Figure 5 shows the simulated fraction of bicarbonate in the groundwater, and the locations of monitoring wells M1–M4. Note the fast path causing lower concentrations and faster breakthrough at ~4.5 m depth. The density of the injected fluid is higher than the resident groundwater, as seen with the higher fraction sinking toward the bottom of the model.

Fluid EC measurements were simulated in the monitoring wells (Figure 6). As the monitoring wells are screened in the saturated zone, each simulated measurement represents the average EC value in the saturated zone.

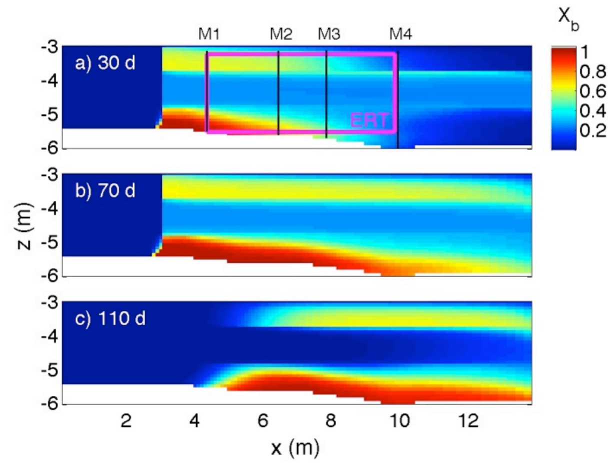

Figure 5. Fraction of injected fluid at (a) 30, (b) 70 and (c) 110 days after the start of injection. The locations of monitoring wells (M1 to M4) and the region corresponding to the idealized ERT data (magenta line) are shown in (a).

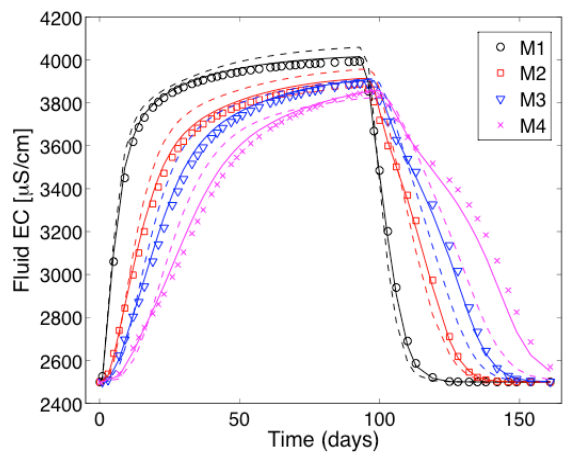

Figure 6. Fluid EC data at monitoring wells M1 to M4. The (synthetic) data are shown with symbols while inversion results are shown with solid and dashed lines for the inversion cases with EC and ERT data smoothed at 0.5 m and 1.1 m, respectively.

Intentionally idealized time-lapse ERT data were "simulated" for cases with increasing distortion to mimic the smoothing that is expected in ERT tomography (Figure 5 shows the region that is considered to contain ERT data). The images in the left column of Figure 7 represent the most ideal case in which ERT tomography images perfectly reflect the spatial variations in electrical conductivity resulting from the injected fluid.

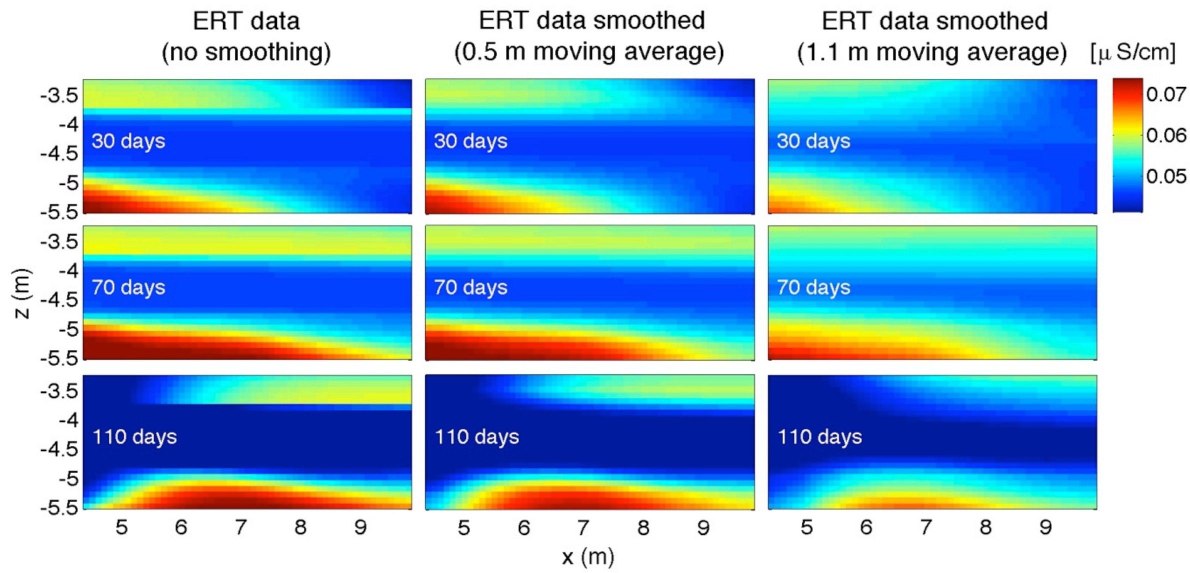

Figure 7. Intentionally idealized (synthetic) time-lapse ERT data. The images in the left column represent the most ideal case, the true distribution (i.e., the ERT tomography images perfectly reflect the spatial variations in electrical conductivity resulting from the injected fluid) at three times. The middle and right columns are increasingly smoothed versions of the true distribution obtained with a moving average of 0.5 m (middle column) and 1.1 m (right column), respectively; they are also considered idealized in the sense that real ERT tomography results would undoubtedly be more corrupted with imaging artifacts and affected by measurement noise. The quantity shown is the electrical conductivity [μS/cm], defined as the inverse of electrical resistivity.

The middle and right columns are increasingly smoothed versions of the true distribution obtained with a moving average of 0.5 m (middle column) and 1.1 m (right column), respectively. While they represent tomograms with error having been introduced in the tomography procedure, in actuality, these cases are still considered idealized in the sense that real ERT tomography results would undoubtedly be more corrupted with imaging artifacts and affected by measurement noise.

Figure 8 shows results for hydrogeophysical inversion using the EC data (Figure 6) and the idealized ERT data (Figure 7) together. With a pixel-based (or layered) parameterization, a total of 9 unknown parameters are estimated, including 8 log permeability values for different depth ranges and also a parameter from the petrophysical model (the cementation exponent *m* of Archie's law). With the most ideal case in which the ERT map perfectly represents the true distribution of the electrically conductive tracer, the parameter estimates are correct. But when including the ERT data with a small amount of error (middle column of Figure 7), whether starting the inversion with the unknown parameters set to the true values or set to initial guesses of -10.75, the parameter estimates are biased (see Figures 8a and 8b, respectively), having moved away from the true values. When more significant smoothing is present in the ERT data (right column in Figure 7), the parameter estimates become even more inaccurate (see Figure 8).

These results highlight the importance of choosing a proper method for including geophysical data in a hydrogeophysical inversion. When implementing tomograms directly as calibration data, even for idealistic cases, errors inherent to geophysical tomography can introduce bias or inaccuracy into the hydrological parameter estimates. However, as has been demonstrated (Commer et al., 2012), when using the approach in which the geophysical measurements themselves are simulated directly in the inversion, such bias is overcome and improved results are obtained.

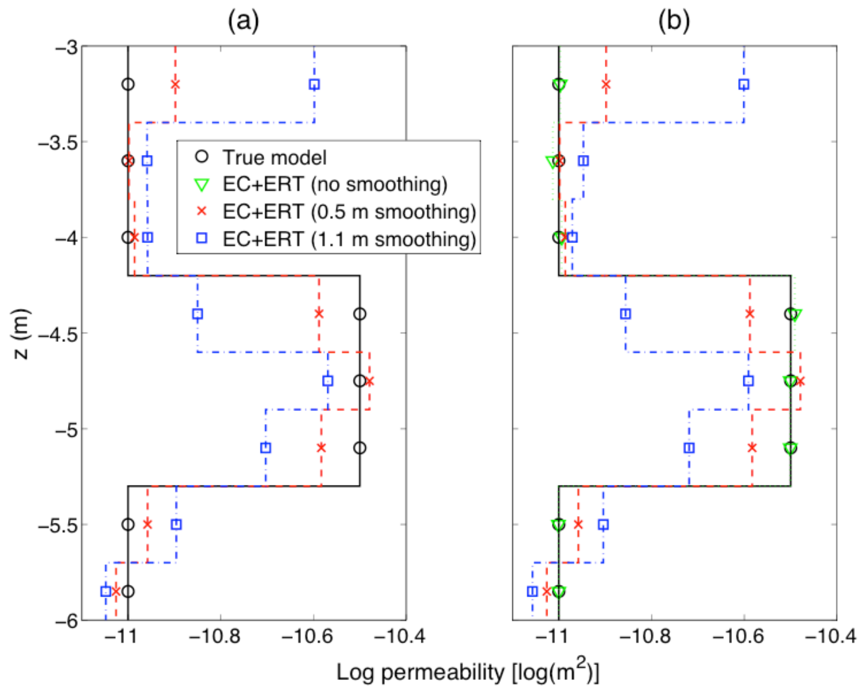

Figure 8. True and estimated log permeability values when starting with the initial guesses (a) set to the true values and (b) set uniformly to -10.75. The *m* value of Archie's law was also estimated.

CONCLUSIONS

One of the main difficulties in building biogeochemical models for complex field experiments continues to stem from uncertainty in the basic heterogeneous hydrological properties, such as permeability and porosity. Thus, testing, improving, and refining techniques for estimating such properties remains an essential research topic in hydrogeology.

We presented examples intended to highlight the importance of decisions made in the process of parameterizing heterogeneity and incorporating data into a hydrogeophysical inversion. Some observations of interest are as follows: When incorporating spatial data, like geometry inferred from core descriptions or geophysical data, uncertainty in such data must be accounted for, and parameterization in regions without data coverage must be done carefully, otherwise biased estimates of hydrological parameters may result. In a geostatistical parameterization, a variety of considerations are necessary, such as the coverage and spacing of pilot points.

Incorporating time-lapse geophysical monitoring data has long been viewed as a promising way to supplement limited hydrological measurements with more spatially extensive and cost effect information. But the manner in which geophysical data are introduced into an inverse modeling framework determines whether they will in fact improve hydrological parameter estimates or introduce bias. Calibrating a hydrological model with a geophysical image that contains errors from tomography (smoothing and streaks) may be less effective than directly simulating the geophysical measurements in a coupled fashion within the inversion, as has been demonstrated here and by others (Hinnell et al., 2010; Commer et al., 2012).

ACKNOWLEDGMENT

This material is based on work supported by the Subsurface Science Scientific Focus Area funded by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research, under award DE-AC02-05CH11231.

REFERENCES

Anderson, R.T., et al. (2003), Stimulating the in situ activity of Geobacter species to remove uranium from the groundwater of a uranium-contaminated aquifer, *Appl. Environ. Microbiol.*, *69*(10), 5884–5891.

Archie G.E. (1942), The electrical resistivity log as an aid in determining some reservoir characteristics, Trans. Am. Inst. Min. Metall. Pet. Eng., 146, 54–62.

Commer, M., M.B. Kowalsky, S. Finsterle, and G.A. Newman (2012), Advances in hydrogeophysical joint inversion, TOUGH Symposium, Sept. 17–19.

Deutsch, C.V., and A.G. Journel (1992), GSLIB: Geostatistical Software Library and User's Guide, Oxford Univ. Press, New York.

Finsterle, S. (2004), Multiphase inverse modeling: Review and iTOUGH2 applications, *Vadose Zone J.*, *3*, 747–762.

Hinnell, A.C., T.P.A. Ferré, J. A. Vrugt, J.A. Huisman, S. Moysey, J. Rings, and M.B. Kowalsky (2010), Improved extraction of hydrologic information from geophysical data through coupled hydrogeophysical inversion, *Water Resour. Res.*, 46, W00D40, doi:10.1029/2008WR007060.

Kowalsky, M.B., S. Finsterle, and Y. Rubin (2004), Estimating flow parameter distributions using ground-penetrating radar and hydrological measurements during transient flow in the vadose zone, *Adv. Water Resour.*, *27*(6), 583–599.

Kowalsky, M.B., S. Finsterle, J. Peterson, S. Hubbard, Y. Rubin, E. Majer, A. Ward, and G. Gee (2005), Estimation of field-scale soil hydraulic and dielectric parameters through joint inversion of GPR and hydrological data, *Water Resour. Res.*, 41, W11425, doi:10.1029/2005WR004237.

Kowalsky, M.B., J. Birkholzer, J. Peterson, S. Finsterle, S. Mukhopadhyay, and Y. Tsang (2008), Sensitivity analysis for joint inversion of GPR and thermal-hydrological data from a large-scale underground heater test, *Nuclear Technology*, *164*(2), 196-206.

Kowalsky, M.B., S. Nakagawa, and G.J. Moridis (2010), Feasibility of monitoring gas hydrate production with time-lapse vertical seismic profiling, *SPE J.*, 15(3), 634-645. SPE-132508-PA, doi: 10.2118/132508-PA.

Kowalsky, M.B., E. Gasperikova, S. Finsterle, D. Watson, G. Baker, and S.S. Hubbard (2011), Coupled modeling of hydrogeochemical and electrical resistivity data for exploring the impact of recharge on subsurface contamination, *Water Resour. Res.*, *47*, W02509, doi:10.1029/2009WR008947.

Kowalsky, M.B., S. Finsterle, K.H. Williams, C. Murray, M. Commer, D. Newcomer, A. Englert, C.I. Steefel, and S.S. Hubbard (2012), On parameterization of the inverse problem for estimating aquifer properties using tracer data, *Water Resour. Res.*, *48*, W06535, doi:10.1029/2011WR011203.

Moore, C., and J. Doherty (2006), The cost of uniqueness in groundwater model calibration, *Adv. Water Resour.*, *29*(4), 605–623.

Pruess, K., C. Oldenburg, and G. Moridis, *TOUGH2 User's Guide, Version 2.0*, Report LBNL-43134, Lawrence Berkeley National Laboratory, Berkeley, Calif., 1999.

Vrionis H.A., R.T. Anderson, I. Ortiz-Bernad, K.R. O'Neill, C.T. Resch, A.D. Peacock, Dayvault, D.C. White, P.E. Long, D.R. Lovley (2005), Microbiological and geochemical heterogeneity in an in situ uranium bioremediation field site, *Appl. Environ. Microbiol,* 71:6308–6318.

Williams, K.H., et al. (2011), Acetate availability and its influence on sustainable bioremediation of uranium-contaminated groundwater, *Geomicrobiol. J.*, *28*(5–6), 519–539, doi:10.1080/01490451.2010.520074:2011.

Yabusaki, S. B., et al. (2007), Uranium removal from groundwater via in situ biostimulation: Field-scale modeling of transport and biological processes, *J. Contam. Hydrol.*, 93, 216–235.